

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2005, 12:11:19 ; Search time 41 Seconds
(without alignments)
624.236 Million cell updates/sec

Title: US-10-006-867-2

Perfect score: 1392
Sequence: 1 MMWFOQLSFLPSALVWTS.....YDTAPCPINNERTRLLSRDI 266

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	361.5	26.0	246	2 T13747	hypothetical prote
2	288	20.7	271	2 T19654	hypothetical prote
3	120	8.6	547	2 C88456	protein W03A5.2 [i
4	103	7.4	332	2 E69312	nitrate reductase
5	98.5	7.1	1120	2 A10561	integral membrane
6	97	7.0	604	2 E72523	carbon starvation
7	96	6.9	495	2 S43882	NADH2 dehydrogenas
8	94	6.8	495	1 S16447	NADH2 dehydrogenas
9	93.5	6.7	379	2 D96506	hypothetical prote
10	93.5	6.7	429	2 F71651	putrescine-ornithi
11	93	6.7	1120	2 F90693	mechanosensitive c
12	93	6.7	1120	2 B85544	mechanosensitive c
13	93	6.7	1120	2 H64733	probable membrane
14	92.5	6.6	473	2 AE0801	probable amino aci
15	91.5	6.6	373	2 T24487	hypothetical prote
16	91.5	6.6	712	2 T02479	potassium transpor
17	90.5	6.5	614	2 E86194	hypothetical prote
18	90	6.5	400	2 C82503	hypothetical prote
19	90	6.5	492	2 S59107	NADH2 dehydrogenas
20	90	6.5	528	2 S19366	hypothetical prote
21	89.5	6.4	440	2 B64090	dicarboxylate tran
22	89.5	6.4	443	2 AE1706	hypothetical prote
23	89.5	6.4	495	2 G71019	hypothetical prote
24	89.5	6.4	599	2 T24333	hypothetical prote
25	89	6.4	428	2 A10034	probable transmemb
26	89	6.4	757	2 T42593	hypothetical prote
27	88.5	6.4	430	2 E75217	transporter PAB217
28	88.5	6.4	844	2 T23656	hypothetical prote
29	88	6.3	1165	2 A46180	adenylyl cyclase t

RESULT 1
T13747
hypothetical protein 22E5.9 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13747
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17668
A:Accession: T13747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-246 <MUR>
A:Cross-references: UNIPROT:O77262; EMBL:AL031765; NID:el371523; PID:el326055; PIDN:CAA
C:Genetics:
A:Cross-references: FlyBase:FBgn0000667
A:Introns: 22/2; 63/1; 192/2
A>Note: EG:22E5.9
C:Superfamily: Caenorhabditis elegans hypothetical protein C33A11.2

Query Match 26.0%; Score 361.5; DB 2; Length 246;
Best Local Similarity 33.6%; Pred. No. 3.5e-24;
Matches 78; Conservative 50; Mismatches 91; Indels 13; Gaps 4;

QY 11 LPSALVWITSAAFIYSYITAVTLRHIDPALPYISDTGTVAPEKCLEGAMNIAAVLCIAT 70
DB 7 LPVLTLFIQVTFLTGTYPFAVLEGHVVPVPIYSDAATYSPSCVFGQLINIGSVLIGIT 66
QY 71 IVRYRKQVHAL--SPE-ENVIIKLNKAGLVGLISCLGISIVANFQKTTLFAHVSGAV 126
DB 67 IVYRQVQLYEHHPDLGSLVLRQNLALWFLGSLGISFVGNFQETNVRIVHIGAF 126
QY 127 LTFMGSLYMFQVTLISYQMPKHGKQVFWIRLLVIMCVGSALSML---TCSSVLHSG 183
DB 127 CFCGCGTLYFWNQALISYILFPMSGTRINAHRLGMSVVCTILFILLAVTGVMSHLFKG 186
QY 184 NFGTDLEQKLHNPDKGVLHMITTAENWSMSPFFGFFLYIRDFOKISL 235
DB 187 -----QNPWKYPDSGGWYFHVSSISEWVIATVFPFILSFTNFRDVS 232

RESULT 2
T19654
hypothetical protein C33A11.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19654
R:Gajadaty, S.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19158
A:Accession: T19654

hemolysin III (ypl
GGDEF family prote
metabolite transpo
hypothetical prote
bicyclomycin resis
bicyclomycin resis
probable polyacch
glutathione-regula
transporter homolo
glycerophosphoryl
hypothetical prote
hypothetical prote
bicyclomycin resis
P22L4.12 protein -
epermidine/putresc


```
Best Local Similarity 24.1%; Pred. No. 1.7;
Matches 54; Conservative 34; Mismatches 84; Indels 52; Gaps 11;

QY 8 LSPALVIVTSAAFISYITAVTLHHIDPALFYISDTGTVAPEKCLFGAM--LNI--- 62
Db 277 LCFTPT---FIYTLGAIAIYTSLSLTQAQID--LKKIAYSSVAHMLNVIGHFSLNIQI 331
QY 63 -----AAVLCIATIVRYKQ-----VHALSPENVIKLNKAGLV 98
Db 332 GGSILLMLSHGLVSALFLCVGLYDRHKTRLVRYGGVLVSTMPNFSTIFFEFTLANMSL 391
QY 99 -GIISCLG--LSIVANFQKTLFAAHVSGAVLTFGMSLYMFVQTILSYQMOP----- 148
Db 392 PGTSSPFGFLILVGAQFQNSLVATLRALGMILGAAYSLWLY-NRVVSNLAPDFLYKFS 450
QY 149 KINGKQVFWIR--LLLVIMCGVSALSMITC-----SSVLHSGNF 185
Db 451 DLNGREVFIFLPFLVGVMVGVYKPVFLDMHTSVSNLVQHGKF 494

RESULT 9
D96506
hypothetical protein T12C22.2 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D96506
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
aneen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: D96506
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-379 <STO>
A/Cross-references: UNIPROT:Q9LPF6; GB:AE005173; NID:98655985; PIDN:AAF78258.1; GSPDB:GN
C/Genetics:
A/Map position: 1
C/Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 6.7%; Score 93.5; DB 2; Length 379;
Best Local Similarity 21.8%; Pred. No. 1.4;
Matches 47; Conservative 53; Mismatches 95; Indels 21; Gaps 9;

QY 10 FLPSALVIVTSAAFISYITAVTLHHIDPALFYISDTGTVAPEKCLFG-AMLNIAA---- 64
Db 90 YIP--LULLPSSASVSESSCSLKYI--VLIVLGVIIAGDNMLYSVGLLYLSASTYS 145
QY 65 VLCIATIVRYKQVHALSPENVIKLNKAGLVGLISCLGISIVANFQKTLFAAHVSG 124
Db 146 LICATQLAFNAFVSYPINAKETALINLSWLLSFAALIALNDADTPSGVSKYIVG 205
QY 125 AVLTFGMSLYMFVQVILSYQMOPKINGQVFWIRLLLVIMCGVSALSMITCSSLVHSGN 184
Db 206 FVCTLAASALYSLLSLMQPFSE-KILKRETPSVVLEMQIYTSLSVA-----TCVSVI--GL 258
QY 185 EGTDLQKLNWPE--DKG---YVLHMITTAEMSM 215
Db 259 FASGEWRTLHGMEGYHKGQASYVLTAVTAVTQV 294

RESULT 10
F71651
putrescine-ornithine antiporter (pote) RP483 - Rickettsia prowazekii
C/Species: Rickettsia prowazekii
C/Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
```

```
C/Accession: F71651
R/Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A/Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A/Reference number: A71630; MUID:99039499; PMID:9823893
A/Accession: F71651
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-429 <AND>
A/Cross-references: UNIPROT:Q9ZD63; GB:AJ235272; GB:AJ235269; NID:93861033; PIDN:CAA1493
A/Experimental source: strain Madrid E
C/Genetics:
A/Gene: pote; RP483
C/Superfamily: L-lysine transport protein

Query Match 6.7%; Score 93.5; DB 2; Length 429;
Best Local Similarity 24.1%; Pred. No. 1.6;
Matches 48; Conservative 37; Mismatches 79; Indels 35; Gaps 10;

QY 11 LPSALVIVTSAAFISYITAVTLHHIDPA-----LPYISDTGTVAPEKCLFGAMLN-- 61
Db 217 IPRAIIIGTCVAFYIINSIGIILIPASELINSKAPY-ADAAATL-----LFGGTWSKV 270
QY 62 ---IAVLCIATI---YVRYKQVHALSPENVIKLN-----NKAGLVGLISCLGISIV 109
Db 271 ITVIASVICIGTLNAWVLTSQIALGLAEDGLLPKFFAKNSNNAPYGIISCLGITPL 330
QY 110 ANFQKTLFAAHVSGAVLTFGMSLYMFVQVILSYQMOPKI-HGQVFWIRLLLVIMCGV 168
Db 331 LLFTSNNFQKQIT-QIIDFSV-IAFLFVYLICSLAPLKFVSSKENFSYYLFI-----V 393
QY 169 SALSMLTCSVSLHSGNFGT 187
Db 384 AIISIIIFCTWVYKTPET 402

RESULT 11
P90693
mechanosensitive channel protein [imported] - Escherichia coli (strain O157:H7, substrai
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: P90693
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: P90693
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1120 <HAV>
A/Cross-references: UNIPROT:Q8XD54; GB:BA000007; PIDN:BA833941.1; PID:gl3359975; GSPDB:C
A/Experimental source: strain O157:H7, substrain RIMD 050952
C/Genetics:
A/Gene: ECs0518

Query Match 6.7%; Score 93; DB 2; Length 1120;
Best Local Similarity 19.6%; Pred. No. 5;
Matches 46; Conservative 38; Mismatches 85; Indels 66; Gaps 8;

QY 1 MWVFOGSLFPLSALVIVTSAAFISYITAVTLHHIDPALFYISDTGTVAPEKCLFGAML 60
Db 805 MFW-----AIWSDLITVFSYLDSTLWHYNG-----TEAGAAVVKVNTWGSLL 847
QY 61 -----NIAAVLCIATIVRYKQVHALSPENVIKLNKAGLVGLIS- 102
Db 848 FAITASVVAWALIRNLPGLLEVLVLSRLNRQSGAYAITILNYIIIAVGAMTVFSGSLV 907
QY 103 -----CLGLSIVANFQKTLFAAHVSGAVLTF-----GMGSLYMF 137
Db 908 SWDKLOWLAALSVGLGFLQEIFGNFVSGLIILFERPVRIGDVTVIGTSFGSTVKIR 967
QY 138 VQTILSYQMOPKINGQVFWIRLLLVIMCGVSALSMITCSSLVHSG-NFGTDLEQ 191
```

```

F,892-908/Domain: transmembrane #status predicted <TM9>

Query Match      6.7%  Score 93;  DB 2;  Length 1120;
Best Local Similarity 19.6%;  Pred. No. 5;
Matches 46;  Conservative 38;  Mismatches 85;  Indels 66;  Gaps 8

Qy      1  MMWFQQGLSLPSALVLTWTSAAPIFFSYITATVLLHHIDPALPYISDTGTVAPEKCLFGAWL 60
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
Db      805  MFW-----ALWSDLIIVFSDSDITLWHYNG-----TEAGNAVKNVTMGSL 847

Qy      61  -----NIAAVLCIATYVRYKOVHALSPSEENVIIKLNKAGLVGLIIS- 102
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
Db      848  FAIIASVVAWALIRNLPGLLEVLVLSRLNMQGASYAITTLNIIYIIAVGAMTVFSGLVG 907

Qy      103  -----CLGLSTVANFQKTLTFAAHVSGAVLTF-----GMSLSYMF 137
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
Db      908  SWDKLWLAALSVGLFGLOBIQFNVFSGUIILFERPVRIGDVTWTCGSPSGTWSKIR 967

Qy      138  VQTILSYQMOPKIHGKQVPIRLLLVTCGVSALSMLTCSVLHSG-NFGTDLEQ 191
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
Db      968  ATTIPTDPRKEVIIPNKAF-VTERLINW----SLTDTTIRLIVIRLGVAYGSDLEK 1017

```

RESULT 14
AE0801
probable amino acid transporter SRY2589 [imported] - *Salmonella enterica subsp. C*; Species: *Salmonella enterica subsp. enterica* serovar Typhi
A/Note: this species has also been called *Salmonella typhi*
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AE0801
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.;
th, T.; Comerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.
S.; Moulle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Ste
A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enter*
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AE0801
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-473 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD07591.1; PID:g16503583; GSPDB:GN00176
C/Genetics:
A/Gene: SRY2589

```

Query Match          6.6%; Score 92.5; DB 2; Length 473;
Best Local Similarity 22.8%; Pred. No. 2.1;
Matches 59; Conservative 36; Mismatches 95; Indels 69; Gaps 14;

Qy 13 SALVIWTSAAPIESYTAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNTAIVLCIATYI 72
   ||::: ||| |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 37 SAMLLWIAAFCLFFPLAMVCGELSTGWP--KGGIFWIKAEAFGKRIAWIVVVCF---- 90

Qy 73 VRYKQVHALSPENNVIIKLNKAGVLGLISC-LGLSIVANFQKTLTFAAHVSGAVLT--- 128
   ||:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 91 -----LFSCVLFFPLMLQFGFTAL--GYMIGGGLAENK 121

Qy 129 --FGMGSLYMF-VOTILSYQMOKIHGQVFWIRLL--LVITWCGV---SALSMLTCSVL 180
   ||:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 122 AFIGIGSAVIFWLLTLMN-----IRGWE--WTKIINSIAWCGVFTPSAILILLAVVWL 173

Qy 181 HSGN-FGTDLEQKLHNWPKDGKGYVLHMITTAEWMSMSPFPGFELTYI-----RDF 230
   ||:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 174 CTGHQMCTDYTTAKNMP-DLG---HWDITIVFLSSMMFAFAGLEVAPMIAGRTRNPQRDF 229

Qy 231 QKISLRVEANLHGLTYDT 249
   : |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 230 PRAMAVSAAVIGIYMGVT 248

RESULT 15
T24487
hypothetical protein T05AL1 - Caenorhabditis elegans

```

Query Match	6.6%;	Score 91.5;	DB 2;	Length 373;
Best Local Similarity	23.0%;	Pred. No. 2;		
Matches	50;	Conservative 34;	Mismatches 82;	Indels 51; Gaps 11;
QY	69	ATIVRYKQVHALSPENVIKLNKAGLVGLISCLGLSIVANFOKTTLFAAH-----	121	
Db	46	AVLYITMKH-RQLQTVQNIIFILNCSNVLMCLTSLPTIFITNVTKQMFSPFVKCIPL	104	
QY	122	VSGA---VLTPGMSGLYMFVOTLSYQOMPKHGKQVFIIRLLVWIGVSAISMILTCSS	178	
Db	105	VOGASIFVSTFSLSAIALDYNLVWFRPHKQKLSRSSAMMIAL--IW---VISVVVCM	158	
QY	179	VLHSGNFOTDLE-OKL-----HWNPED--KGYVLHMITTAAEWSMSPFGFFLT	235	
Db	159	-----YGYWMDVEKLNGLCGEYCSHHWPLAEVRKGYTFLVLT-----QFLPFPATMAFC	208	
QY	226	YIRDFOKISLVEANLHGLTLYDTAPCPINNERTLL	262	
Db	209	YNYIFSLRQRVETKKL-----SRSQILL	234	

Search completed: January 4, 2005, 12:20:57
Job time : 43 secs